(6) SOME INEQUALITIES ON CHARACTERISTIC ROOTS OF MATRICES*

By

(8) T. W. Anderson and S. Das Gupta

Columbia University (5) 205 500

Department of Mathematical Statistics

(9) April 1, 1963

(12) *This research was sponsored by the Office of Naval Research under Contract Number Nonr-266(33), Project Number NR 042-034. Reproduction in whole or in part is permitted for any purpose of the United States Government.

1. SUMMARY AND INTRODUCTION

Some inequalities on the characteristic roots of positive definite matrices and of matrices derived from them are obtained. Some of these are generalizations of results of Roy (1954) which give upper and lower bounds on all the roots of the product of two positive definite matrices in terms of the maximum and the minimum roots of the factor matrices. These various results are found to be useful in the study of statistical problems concerning the normal multivariate distribution, and the authors hope that they might be useful in other mathematical statistics investigations also.

2. BOUNDS ON CHARACTERISTIC ROOTS

The results are based on the following theorem from Courant and Hilbert (1953, p. 31):

Theorem 2.1 *For any symmetric matrix* A *of order* p *and for any* $i-1$ $(1 \leq i \leq p)$ *column vectors* α_j $(j = 1, \ldots, i-1)$

$$\max_{\substack{x' \alpha_j = 0 \\ j = 1, \ldots, i-1}} \frac{x' A x}{x' x} \geq ch_i(A) . \tag{2.1}$$

Here $ch_1(A) \geq ch_2(A) \geq \ldots \geq ch_p(A)$ are the characteristic roots of A.

<u>Lemma 2.1</u> <u>Let</u> A <u>be a symmetric matrix of order</u> p, <u>and</u> B <u>be a positive definite matrix of order</u> p. <u>Then for any</u> $i-1$ $(1 \leq i \leq p)$ <u>column vectors</u> α_j $(j = 1, \ldots, i-1)$

$$\max_{\substack{x'\alpha_j = 0 \\ j = 1, \ldots, i-1}} \frac{x'Ax}{x'Bx} \geq ch_i\,(AB^{-1}) \;. \tag{2.2}$$

<u>Proof</u>. There exists a nonsingular matrix T such that $B = T'T$. Then (2.2) is identical to

$$\max_{\substack{y'\beta_j = 0 \\ j = 1, \ldots, i-1}} \frac{y'Ey}{y'y} \geq ch_i\,(E) \;, \tag{2.3}$$

where $y = Tx$, $E = (T')^{-1} \; AT^{-1}$, $\beta_j = (T')^{-1}\alpha_j$, and $ch_i\,(AB^{-1}) = ch_i\,[AT^{-1}\,(T')^{-1}] = ch_i\,(E)$. The lemma now follows from Theorem 2.1 and the fact that (2.3) is equivalent to (2.1).

<u>Theorem 2.2</u> <u>Let</u> A <u>be a symmetric matrix of order</u> p, <u>and let</u> B <u>and</u> C <u>be positive definite matrices of order</u> p. <u>Then for any</u> i $(1 \leq i \leq p)$

$$ch_i\,(AB^{-1}) \leq ch_j\,(AC^{-1}) \;\; ch_k\,(CB^{-1}) \tag{2.4}$$

and

$$\mathrm{ch}_{p-i+1}(AB^{-1}) \geq \mathrm{ch}_{p-j+1}(AC^{-1})\, \mathrm{ch}_{p-k+1}(CB^{-1}) \tag{2.5}$$

for $j + k \leq i + 1$.

Proof. Let $\alpha_1, \ldots, \alpha_{j-1}$ be the first $j-1$ characteristic vectors of A in the metric of C, that is, $A\alpha_h = \mathrm{ch}_h(AC^{-1})\, C\alpha_h$, and let $\alpha_j, \ldots, \alpha_{j+k-2}$ be the first $k-1$ characteristic vectors of C in the metric of B; that is, $C\alpha_{j+\ell} = \mathrm{ch}_{\ell+1}(CB^{-1})\, B\alpha_{j+\ell}$. Then

$$\begin{aligned}
\mathrm{ch}_i(AB^{-1}) &\leq \mathrm{ch}_{j+k-1}(AB^{-1}) \\
&\leq \max_{\substack{x'\alpha_h = 0 \\ h = 1, \ldots, j+k-2}} \frac{x'Ax}{x'Bx} \\
&= \max_{\substack{x'\alpha_h = 0 \\ h = 1, \ldots, j+k-2}} \frac{x'Ax}{x'Cx} \cdot \frac{x'Cx}{x'Bx} \\
&\leq \max_{\substack{x'\alpha_h = 0 \\ h = 1, \ldots, j+k-2}} \left[\frac{x'Ax}{x'Cx}\right] \cdot \max_{\substack{x'\alpha_h = 0 \\ h = 1, \ldots, j+k-2}} \left[\frac{x'Cx}{x'Bx}\right] \\
&\leq \max_{\substack{x'\alpha_h = 0 \\ h = 1, \ldots, j-1}} \left[\frac{x'Ax}{x'Cx}\right] \cdot \max_{\substack{x'\alpha_h = 0 \\ h = j, \ldots, j+k-2}} \left[\frac{x'Cx}{x'Bx}\right]
\end{aligned}$$

$$= \quad ch_j\,(AC^{-1})\; ch_k\,(CB^{-1})\ .$$

To derive (2.5) , we replace A by $-A$ in (2.4) to get

$$ch_i\,(-AB^{-1}) \le ch_j\,(-AC^{-1})\; ch_k\,(CB^{-1}) \tag{2.6}$$

for $j + k \le i + 1$. Note that for any symmetric matrix T of order p

$$ch_i\,(-T) = -\,ch_{p-i+1}\,(T)\ , \quad (i = 1, \ldots, p)\ , \tag{2.7}$$

Now (2.5) follows from (2.6) with use of the relation (2.7) .

Corollary 2.2.1 If A is a symmetric matrix of order p , and B is a positive definite matrix of order p, then for any $i\ (1 \le i \le p)$

$$ch_p\,(B)\; ch_i\,(A) \le ch_i\,(AB) \le ch_i\,(A)\; ch_1\,(B)\ , \tag{2.8}$$

and furthermore, if A is positive definite , then

$$\frac{ch_i^2\,(AB)}{ch_1\,(A)\; ch_1\,(B)} \le ch_i\,(A)\; ch_i\,(B) \le \frac{ch_i^2\,(AB)}{ch_p\,(A)\; ch_p\,(B)}\ . \tag{2.9}$$

Proof. If we write $C = I$ and replace B by B^{-1} in Theorem 2.2 , we obtain

$$ch_i\,(AB) \le ch_j\,(A)\; ch_k\,(B) \tag{2.10}$$

and

$$\mathrm{ch}_{p-i+1}(AB) \geq \mathrm{ch}_{p-j+1}(A)\,\mathrm{ch}_{p-k+1}(B) \tag{2.11}$$

for $j + k \leq i + 1$. By letting $j = i$ in (2.10), replacing $p - i + 1$ and $p - j + 1$ by i in (2.11), and letting $k = 1$ in both, we obtain (2.8). When A is also positive definite, then interchanging A and B in (2.8), we obtain

$$\mathrm{ch}_p(A)\,\mathrm{ch}_i(B) \leq \mathrm{ch}_i(AB) \leq \mathrm{ch}_i(B)\,\mathrm{ch}_1(A). \tag{2.12}$$

Now (2.9) follows from (2.8) and (2.12).

From Corollary 2.2.1, we obtain

$$\mathrm{ch}_p(A)\,\mathrm{ch}_p(B) \leq \mathrm{ch}_i(AB) \leq \mathrm{ch}_1(A)\,\mathrm{ch}_1(B)$$

$(1 \leq i \leq p)$. This result was obtained previously by Roy (1954).

Corollary 2.2.2 *Let* A *be a symmetric matrix of order* p. *Then* $\mathrm{ch}_i(A)$ $(1 \leq i \leq p)$ *is not less than the minimum characteristic root of any principal minor of* A *of order* i.

Proof. It will be enough to show that $\mathrm{ch}_i(A)$ is at least equal to the minimum characteristic root of A_i, where A_i is the submatrix formed by the first i rows and columns of A. Let D be a diagonal matrix with diagonal elements $d_1, \ldots, d_p$ such that $d_j = 1$ for $j \leq i$,

$d_j < 1$ for $j > i$. Then from (2.8)

$$ch_i\ (DAD) = ch_i\ (D^2A) \leq ch_1\ (D^2)\ ch_i\ (A) = ch_i\ (A) .$$

Thus

$$ch_i\ (A) \geq \lim_{d_j \to 0\ (j > i)} ch_i\ (DAD) .$$

Now

$$\lim_{d_j \to 0\ (j > i)} (DAD) = \left[\begin{array}{c|c} A_i & 0 \\ \hline 0 & 0 \end{array}\right] ,$$

where the matrix is partitioned into i and $p-i$ rows and columns.

Hence

$$ch_i\ (A) \geq \lim_{d_j \to 0\ (j > i)} ch_i\ (DAD)$$

$$= ch_i\ (A_i) = \min ch\ (A_i) .$$

<u>Theorem 2.3</u> <u>Let</u> A <u>and</u> B <u>be two symmetric matrices of order</u> p. <u>Then for any</u> i $(1 \leq i \leq p)$

$$ch_i\ (A+B) \leq ch_j\ (A) + ch_k\ (B) \qquad (2.13)$$

and

$$\mathrm{ch}_{p-i+1}(A+B) \geq \mathrm{ch}_{p-j+1}(A) + \mathrm{ch}_{p-k+1}(B) \tag{2.14}$$

for $j + k \leq i + 1$.

Proof. Let $\alpha_1, \ldots, \alpha_{j-1}$ be the characteristic vectors of A corresponding to the $j-1$ largest roots, and $\alpha_j, \ldots, \alpha_{j+k-2}$ be the characteristic vectors of B corresponding to the $k-1$ largest roots . Then

$$
\begin{aligned}
\mathrm{ch}_i(A+B) &\leq \mathrm{ch}_{j+k-1}(A+B) \\
&\leq \max_{\substack{x'\alpha_h = 0 \\ h = 1, \ldots, j+k-2}} \frac{x'(A+B)\,x}{x'x} \\
&\leq \max_{\substack{x'\alpha_h = 0 \\ h = 1, \ldots, j+k-2}} \left[\frac{x'Ax}{x'x}\right] + \max_{\substack{x'\alpha_h = 0 \\ h = 1, \ldots, j+k-2}} \left[\frac{x'Bx}{x'x}\right] \\
&\leq \max_{\substack{x'\alpha_h = 0 \\ h = 1, \ldots, j-1}} \left[\frac{x'Ax}{x'x}\right] + \max_{\substack{x'\alpha_h = 0 \\ h = j, \ldots, j+k-2}} \left[\frac{x'Bx}{x'x}\right] \\
&= \mathrm{ch}_j(A) + \mathrm{ch}_k(B) .
\end{aligned}
$$

The inequality (2.14) is obtained from (2.13) by replacing A and B by $-A$ and $-B$, respectively.

It follows from Theorem 2.3 that, if $C-A$ is a positive semi-definite matrix of order p, then $ch_i(C) \geq ch_i(A)$, $i = 1, \ldots, p$ [as given by Courant and Hilbert (1953, p. 33)].

If A and B have some common characteristic vectors, then the bounds for $ch_i(AB)$ and $ch_{p-i+1}(AB)$ in (2.10) and (2.11) and the bounds for $ch_i(A+B)$ and $ch_{p-i+1}(A+B)$ in (2.13) and (2.14), respectively, can be improved. In particular, if A and B have the same characteristic vectors, then $ch_i(AB) = ch_i(A)\, ch_i(B)$, $ch_i(A+B) = ch_i(A) + ch_i(B)$. A condition for this is $AB = BA$.

REFERENCES

Courant, R., and Hilbert, D. (1953). Methods of Mathematical Physics. Vol. 1, Interscience Publishers Inc., New York

Roy, S. N. (1954). A useful theorem in matrix theory. Proc. Amer. Math. Soc., 5, 635-638.